



**QUEEN'S
UNIVERSITY
BELFAST**

The Coxian phase-type distribution as a contribution to the multilevel model of in-hospital mortality

Zenga, M., Marshall, A. H., Crippa, F., & Mitchell, H. (2016). The Coxian phase-type distribution as a contribution to the multilevel model of in-hospital mortality. *Communications in Statistics: Theory and Methods*, 45(6), 1819-1830. <https://doi.org/10.1080/03610926.2015.1038394>, <https://doi.org/10.1080/03610926.2015.1038394>

Published in:

Communications in Statistics: Theory and Methods

Document Version:

Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

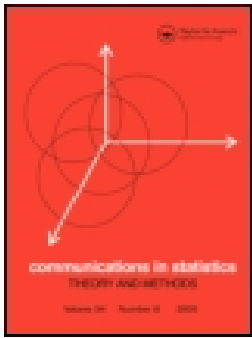
This is an Accepted Manuscript of an article published by Taylor & Francis in Communications in Statistics - Theory and Methods on 09/11/2015, available online: <http://www.tandfonline.com/10.1080/03610926.2015.1038394>

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.



The coxian phase-type distribution as a contribution to the multilevel model of in-hospital mortality

Mariangela Zenga, Adele H. Marshall, Franca Crippa & Hannah Mitchell

To cite this article: Mariangela Zenga, Adele H. Marshall, Franca Crippa & Hannah Mitchell (2015): The coxian phase-type distribution as a contribution to the multilevel model of in-hospital mortality, Communications in Statistics - Theory and Methods, DOI: [10.1080/03610926.2015.1038394](https://doi.org/10.1080/03610926.2015.1038394)

To link to this article: <http://dx.doi.org/10.1080/03610926.2015.1038394>



Accepted author version posted online: 09 Nov 2015.



Submit your article to this journal [↗](#)



Article views: 11



View related articles [↗](#)



View Crossmark data [↗](#)

THE COXIAN PHASE-TYPE DISTRIBUTION AS A CONTRIBUTION TO THE MULTILEVEL MODEL OF IN-HOSPITAL MORTALITY

Mariangela Zenga

University of Milano-Bicocca, Milano, Italy; mariangela.zenga@unimib.it

Adele H. Marshall

Queen's University Belfast, Belfast, Northern Ireland(UK); a.h.marshall@qub.ac.uk

Franca Crippa

University of Milano-Bicocca, Milano, Italy; franca.crippa@unimib.it

Hannah Mitchell

Queen's University Belfast, Belfast, Northern Ireland(UK); h.mitchell3@qub.ac.uk

Key Words: Coxian phase-type distribution, Grouped survival model, multilevel model.

ABSTRACT This paper presents multi-level models that utilise the Coxian phase-type distribution in order to be able to include a survival component in the model. The approach is demonstrated by modelling patient length of stay and in-hospital mortality in geriatric wards in Italy. The multi-level model is used to provide a means of controlling for the existence of possible intra-ward correlations, which may make patients within a hospital more alike in terms of experienced outcome than patients coming from different hospitals, everything else being equal. Within this multi-level model we introduce the use of the Coxian phase-type distribution to create a covariate that represents patient length of stay or stage (of hospital care). Results demonstrate that the use of the multi-level model for representing the in-patient mortality is successful and further enhanced by the inclusion of the Coxian phase-type distribution variable (stage covariate).

1. INTRODUCTION

The proportion of elderly people in the population continues to increase in all European countries. Older people represent the fastest growing sector of society and account for the largest increase in hospital admissions (OECD, 2004). A growth of health care service dedicated to the over 65 year old is required with increased expenditure. The increase of admission to hospitals of elderly patients with multiple co-morbidities results in an overall increase in patient length of stay (LOS). In addition, geriatric patients are at highest risk of acquired disability, cognitive decline, or admission to residential care, either as a consequence of illness or as an unfortunate consequence of treatment (Ellis *et al.*, 2011). For example, in Italy in 2012 the elderly people comprise 37% of the admissions to hospital consuming nearly half (49%) of the LOS days.

In the last two decades, health care agencies and other stakeholders have placed a lot of emphasis on the measure of and formal definition of hospital quality of care. Measuring and improving hospital quality of care, particularly outcomes of care, is an important focus for clinicians and policy makers. The Institute of Medicine in 1990 defines *quality of health care* as:

...the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge (IOM, 1990)

where, health outcome means the "result, often long term, on the state of the patient well-being, generated by the delivery of health service" (Spiegelhalter, 1998). In general, several authors (for example Donebian, 2005 and DesHarnais *et al.*, 1991) state that the quality of hospital care may need to include multiple indices in order to obtain a valid measure of the important aspects of quality. It is possible then to examine the relationships among various measures or alternatively include one and not another. It is natural to consider the risk-adjustment mechanisms to compare quality across institutions, and in examining practices and cultures in high-performing hospitals. However, a clinical outcome such as LOS

could be considered a proxy of consumption of resources and costs for hospitalizations, while in-hospital mortality could be considered a crucial measure of the quality of care provided, maybe the most important. LOS, though not a full measure of cost, is an indication of resource usage and will be used here as a rough proxy for efficiency. Often, a significant reduction in risk-adjusted LOS over time seems primarily to reflect ongoing financial pressures on hospitals to reduce costs. This also may signify improved ability of hospitals to stabilize patients more quickly, or a trend toward discharging patients earlier and caring for them in outpatient, home, and other non-hospital settings. In a study by Kroch *et al.* (2007), the authors show that declining LOS as well as improved mortality rates reflect discharge of sicker patients resulting in more readmissions whereas clinical outcome such as hospital mortality as the final outcome of treatment in a hospital is considered a crucial measure of the quality of care provided. LOS reverberates, in addition to medical advances and changes in care practice, health policies on the theme of hospitalization, particularly cogent in times of funding restrictions and of marked aging. In this framework, in-hospital mortality rates and LOS have been reckoned as distinct variables and their reciprocal relation largely explored so far (Kroch, Duan, Sillow-Carrol and Meyer, 2007; Lau, Fang and Leung, F. 2013). The aim of this paper is to consider a multi-level model to represent patient in-hospital mortality and incorporate within the multi-level structure, an alternative form of modelling patient's LOS using the Coxian phase-type distribution. Such an approach has never been used before to model survival times within multi-level data. This paper will attempt to assess the feasibility of such an approach for modelling patient's LOS according to different levels of care provision. This will be demonstrated through the analysis of hospital stays in all geriatric wards for all regions in Italy. The focus is not only on LOS but also on the in-hospital mortality rate, whose interrelation has been studied mostly in terms of bivariate analysis (Librero, Peirò and Ordiña, 1999). Our analysis will also pay attention to the difference between hospitals as well as between regions given that the Italian National Healthcare System recently transferred several tasks from the central government to the administration of the regions. The current system gives the regions significant autonomy with revenue

and in organizing services (as healthcare services) designed deliberately to better satisfy the needs of their respective populations. The paper is structured as follows: Section 2 reports the definitions and the descriptions of the methods used to evaluate LOS and in-hospital mortality rate, and Section 3 reports the description of the dataset. Section 4 shows the main results and finally Section 5 presents the conclusions.

2 METHODS

The following section describes the modelling approaches used in this methodology.

2.1 THE LENGTH OF STAY IN GERIATRIC WARDS

The effective care of elderly patients in hospital may be enhanced by accurately modelling the patients' LOS in hospital and the associated costs involved. The management of the care of elderly patients in hospitals may be improved if there was a model to represent and predict the LOS. For instance, if a hospital manager were able to estimate the duration of stay of patients on admission to hospital, the ward could be more efficiently managed with better allocation of beds and resources (Marshall and McClean, 2004). The distribution of the LOS of elderly patients in hospital tends to be highly skewed in nature where there is usually a large peak in the distribution at the start which then gradually tails off as duration increases. Past investigations of modelling LOS in geriatric wards have led to the discovery that a two-term mixed exponential model produces a good representation of patient survival (Millard, 1991). Since then further research has provided evidence that the mixed exponential models can be enhanced with the incorporation of more complex compartmental systems and more sophisticated stochastic models such as the Coxian phase-type distribution.

Coxian phase-type distributions are a subset of the widely used phase-type distributions introduced by Neuts in 1975. They have the benefit of overcoming the problem generality within phase-type distributions by only requiring $2n - 1$ parameters to describe a distribution requiring n phases, whereas the general phase-type distribution requires $n^2 + n$ (Neuts, 1981). Coxian phase-type distributions have been used in a variety of settings from component failure time data and the length of treatment for patients at risk of suicide, to prisoner

remand times and the lifetime of male rats (Faddy, 1994). Marshall et al. (2003) used the Coxian phase-type distribution to model the career progression of students at university where the process can be thought of as a series of transitions through latent phases until an event of leaving the university occurs due to graduation. However most applications of the Coxian phase-type distribution has been in modelling the length of time spent in hospital. In particular Faddy and McClean (2005) showed that the Coxian phase-type distribution was appropriate for describing the length of time United Kingdom geriatric patients spent in care. They have also been used to model the stages of progression of the patients from first entering the hospital through to the individual leaving due to recovery or death. The transitions through the ordered transient states could correspond to the stages in patient care such as diagnosis, assessment, rehabilitation and long-stay care where the patients eventually will then reach the absorbing state of the Coxian phase-type distribution corresponding to their departure from hospital either through discharge, transfer or death (Faddy and McClean, 1999).

Methodologically speaking, the Coxian phase-type distribution represents the time to absorption of a finite latent Markov chain in continuous time where there is a single absorbing state and the stochastic process starts in a transient state. They describe the probability $P(t)$ that the process is still active at time t and differ from the general phase-type distributions in that the transient states (phases) of the model are ordered. The process begins in the first phase and either moves sequentially through the phases or into the absorbing state. In other terms, a Coxian phase-type distribution results when the transient states have a natural order and only forward transitions between them may occur. These phases may be used to describe the stages of a process until termination, in which the number of phases and transition rates are latent and need to be estimated as described in Figure 1.

It is possible to attach to the transient states in this model some real world meaning for example within a hospital environment each of the stages could be thought of as the progression of recovery where the first state could be admittance followed by diagnosis, treatment, rehabilitation. During each state the individual can leave hospital due to discharge, transfer

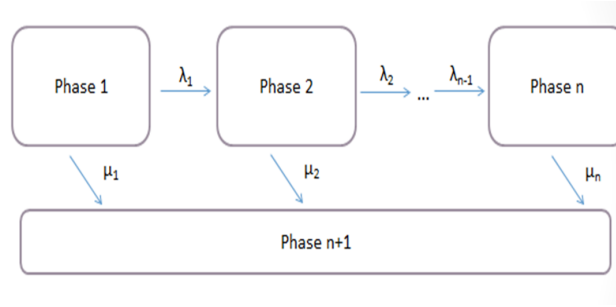


Figure 1: Coxian Phase-type Distribution

or death.

Let $(X(T); t \geq 0)$ be a latent Markov chain in continuous time T with states $1, 2, \dots, n, n+1$ and $X(0) = 1$

for $i = 1, 2, \dots, n-1$

$$\text{prob}\{X(t + \delta t) = i + 1 | X(t) = i\} = \lambda_i \delta t + o(\delta t) \quad (1)$$

and for $i=1,2,\dots,n$

$$\text{prob}\{X(t + \delta t) = n + 1 | X(t) = i\} = \mu_i \delta t + o(\delta t). \quad (2)$$

where states $1,2,\dots,n$ are latent (transient) states of the process and state $n+1$ is the absorbing state, λ_i represents the transition from state i to state $(i+1)$ and μ_i the transition from state i to the absorbing state $(n+1)$.

The probability density function of T can be written as follow:

$$f(t) = \mathbf{p} \exp(\mathbf{Q}t) \mathbf{q} \quad (3)$$

where:

\mathbf{Q} is the matrix of transition rates between states,

$$\mathbf{Q} = \begin{pmatrix} -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & -(\lambda_2 + \mu_2) & \lambda_2 & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -(\lambda_{n-1} + \mu_{n-1}) & \lambda_{n-1} \\ 0 & 0 & 0 & \dots & 0 & -\mu_n \end{pmatrix} \quad (4)$$

$\mathbf{q} = -\mathbf{Q}\mathbf{e} = (\mu_1, \mu_2, \dots, \mu_n)^T$ and $p = (1, 0, \dots, 0)$.

The survival function is given by

$$S(t) = p \exp(\mathbf{Q}t)\mathbf{e} \quad (5)$$

where \mathbf{e} is a column vector of 1's. The number of parameters in a Coxian phase-type distribution is equal to $2n - 1$. The Coxian family is dense in the class of all distributions on $[0, \infty]$ and is appropriate for estimating long-tailed distributions. It turns out that T can be represented as $T = \sum_{k=1}^n Z_k W_k$ with $W_k = T_1 + T_2 + \dots + T_k$ the sum of independent exponential variables, $T_k \sim \exp(\delta_k)$ with hazard rate δ_k , and (Z_1, Z_2, \dots, Z_n) is independent of $\{T_k : 1 \leq k \leq n\}$ with multinomial distribution and probabilities $(\eta_1, \eta_2, \dots, \eta_n)$, $\sum_{k=1}^n \eta_k = 1$. The $S(t) = \sum_{k=1}^n \eta_k S_k(t)$ is a finite mixture of survival distributions, where $S_k(t)$ is the survival distribution of W_k .

Marshall and McClean (2004) derived the formula to represents the LOS in terms of k phases. Let π_i be the proportion of individuals departing the system from the i^{th} phase and calculated as:

$$\begin{aligned} \pi_1 &= \frac{\mu_1}{\lambda_1 + \mu_1} \\ \pi_2 &= \frac{\lambda_1}{\lambda_1 + \mu_1} \cdot \frac{\mu_2}{\lambda_2 + \mu_2} \\ &\dots \\ \pi_k &= \frac{\lambda_1}{\lambda_1 + \mu_1} \cdot \frac{\lambda_2}{\lambda_2 + \mu_2} \cdot \dots \cdot \frac{\lambda_{k-1}}{\lambda_{k-1} + \mu_{k-1}} \end{aligned} \quad (6)$$

Patients may then be divided into groups according to their LOS where the data is grouped in the ratio $\pi_1 : \pi_2 : \dots : \pi_k$. In general the k^{th} LOS group S_k can be determined by the following equation:

$$S_k = \{t_{(j)} : m \sum_{i=1}^{k-1} \pi_i \leq j \leq m \sum_{i=1}^k \pi_i\}, \quad for \quad k = 1, \dots, n \quad (7)$$

where $t_{(1)}, \dots, t_{(m)}$ represents the ordered LOS data for each patient and m represents the number of patients in the data set. The patients' details within each LOS group may then be examined to determine if they have any common characteristics.

2.2 THE ANALYSIS OF IN-HOSPITAL MORTALITY RATE USING A MULTILEVEL REGRESSION MODEL

The use of in-hospital mortality rates in the comparative evaluation of quality of care has been proposed by several authors (for an depth analysis see for example Berta *et al*, 2013). Different statistical methodologies have previously been proposed for risk adjustment of this outcome to account for case-mix differences across healthcare providers so that the outcomes can be legitimately compared despite differences in risk factors. One of the most straightforward approaches to risk adjustment of an outcome, to compare providers, is to estimate an expected value for each provider's outcome based on the relationship between the outcome and its risk factors. The use of multilevel models (also known as random-effects models or hierarchical linear models) was proposed to study the relationships between outcomes and variables related to this phenomenon (Goldstein, 2005; Hox, 2010; Rice and Leyland, 1996). The aim of multilevel models is to control for the existence of a possible intra-ward correlation, which may make patients within a hospital more alike in terms of experienced outcome than patients coming from different hospitals, everything else being equal. Moreover, due to the autonomy in the Italian regional healthcare services, the model has to control the existence of a possible influence that may be made by regional policy for the quality of the geriatric wards. The idea to use multilevel modelling arises from the consideration that, in the Italian national health system, regions have some self-regulation

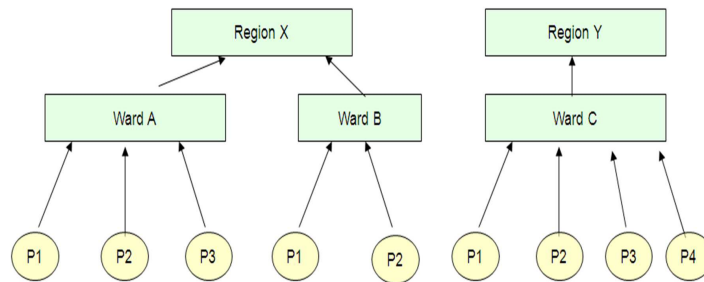


Figure 2: Levels involved in hospitalization outcome

power in health themes and each care unit is allowed to establish its own effectiveness criterion (Fig.2), within a clearly defined legal framework, whilst the central health authority dictates the criteria for hospitalization. Variables for risk adjustment can be measured at each level and the variance of the outcome can be attributed to every level. Multilevel models can provide substantive difference to interpretation and represent a better insight (Jones, Wright and Bell, 2012).

The multilevel model considers the three levels; the patient as the first level (P1, P2, P3 etc.), the ward as the second level and region as the third level, as follows:

$$\eta_{ijk} = \log \frac{\pi_{ijk}}{1 - \pi_{ijk}} = x'_{ijk}\beta + \gamma_{0k} + u_{0jk} + \epsilon_{ijk} \quad (8)$$

where π_{ijk} is the probability of the death for the patient i in ward j in region k , conditional on the variables X ; u_{0jk} is the unobserved hospital random effect of the intercept amongst wards, with $u_{0jk} \sim N(0, \sigma_u^2)$; γ_{0k} is the random variation of the intercept amongst regions,

with $\gamma_{0k} \sim N(0, \sigma_\gamma^2)$ and $\epsilon_{ijk} \sim N(0, \sigma_\epsilon^2)$. Random components at different levels are assumed uncorrelated, whilst non-null correlations are assumed for patients in the same wards or in the same region. The random effect among wards can be interpreted as the relative effectiveness of hospitals with respect to outcome adjusted for fixed coefficients related to patient, ward and regional characteristics. The u_{0jk} estimates showing the specific managerial contribution of the j^{th} wards to the risk-of-mortality and their 95% confidence intervals (CIs) identify hospitals with CIs under or over the mean of the risk-of-mortality.

3. THE ITALIAN GERIATRIC WARDS

The data used in this paper consists of the ordinary admissions of 138,197 patients aged 65 years or older to every geriatric ward in all acute care hospitals (197 geriatric wards in total) operating in the 20 Italian Regions throughout 2009. The data was provided by the Italian Health Care Ministry. Individual Hospital Discharge Charts (HDC) are reported in the dataset including patient information (gender, age, residence etc.), the treatments received during hospitalization including information such as Disease Related Group (DRG), principal and secondary diagnoses and procedures, date of admission and so on.

Patients aged 85 years or older represent 34% of all patients in the dataset. Approximately 42% of patients were male. Seventeen percent of patients were admitted to the department of geriatric medicine through emergency admission. Approximately 53% of patients were admitted for surgery. Chronic patients represented approximately 30% of all patients. Moreover 24% of the admitted patients had a principal diagnosis of circulatory system, 23% as respiratory system problems, 16% problems of the Nervous system. The destination of patients on departure from hospital could be one of several possibilities: the patient may return home; transfer to a nursing home, residential home, another ward, or other hospital; or may die while in hospital. Outcome was coded to describe three locations: home, transfer, or death. Approximately 80% of patients left the geriatric ward to return home, of whom 4% were voluntary discharge; 11% died while in hospital and the remaining 9% transferred.

From a supply point of view, the 197 geriatric wards had on average 24.5 beds in their ward (minimum=4, maximum=147, standard deviation=17.50) and hospitalized on average 701 patients (minimum=58, maximum=7407, standard deviation=960.92) during 2009. The ownership of the hospitals consisted of 83% public, 13% private for-profit, and the remaining 4% private not-for-profit. At a regional level, there were on average 9.8 (SD=7.88) geriatric wards per region, with extremes in the system where *Friuli Venezia Giulia* had only one geriatric ward and *Sicilia* thirty.

4 RESULTS

A multilevel model is used to model patient mortality rates by taking into consideration three levels within the model; the patient, hospital and region. The resulting model was then extended to include a covariate for patient LOS. This covariate is created from fitting the patient LOS to a Coxian phase-type distribution which forms stages of LOS (care in hospital). This covariate is surely endogenous to some extent, being estimated from the same dataset as the model. As a matter of fact, endogeneity is relatively common in multilevel modelling, as it is considered unrealistic that the explanatory variables are always independent of the random error components (Spencer and Fielding, 1999; Ebbes and Bockenholt, 2004). The issue has been addressed explicitly in the case of a covariate non independent of the random effects in the model, indeed this variable may be regarded as endogenous and it could itself be modelled using the same data (Spencer and Fielding, 2002, p.103). In order to adjust the model estimation so to incorporate the extra variance coming from the generated regressor term, we applied the MCMC procedure as the preferred form of estimation for realistically complex models. The key aspect of the approach is that it allows a building block approach to estimation; whereby complex problems are decomposed into lots of small ones which are then linked (Jones *et al.*, 2012). We used the software MLwiN, version 2.27 (Jones and Subramanian, 2012, updated version 2014). The results are outlined below.

4.1 THE LENGTH OF STAY

The highly right skewed patient LOS distribution is illustrated in Figure 3. The average

LOS is eleven days (SD=8.67), the median nine days and the maximum 100 days.

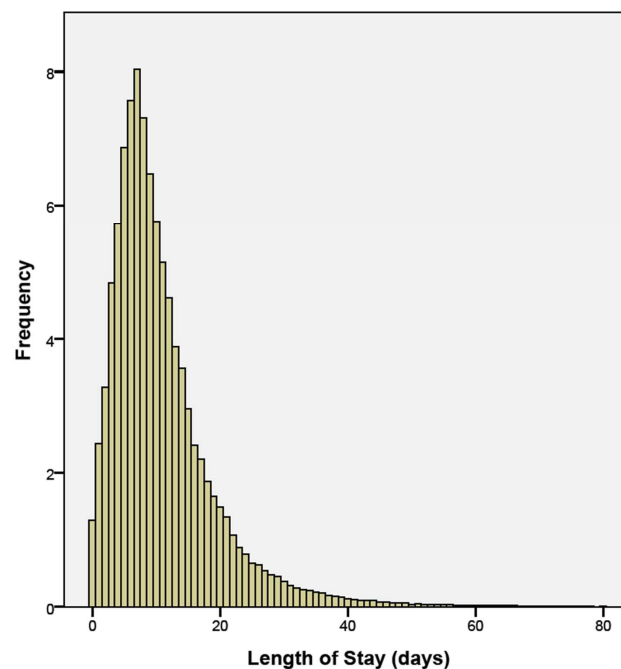


Figure 3: LOS of Geriatric Patients in Italy, 2009

A logrank test was conducted to compare survival distributions for LOS according to the separate patient characteristics. Table 1 shows that, for all variables (except for age), we reject the null hypothesis of equality for the survival function across all categories.

As shown in previous papers, we can suitably represent LOS using the Coxian phase-type distribution by coding the EM-algorithm to perform iteratively in C. At each iteration, the new parameter estimates are calculated by solving a system of homogeneous linear differential equations using the Runge-Kutta method of fourth order. The programme stops when no further significant contribution can be made with the addition of another phase. The Akaike information criterion (AIC) was then calculated and used to find the most appropriate model to represent the data. The results show that the five phase Coxian distribution most suitably represents LOS of geriatric patients in Italy. The loglikelihood for 5 phases was -458860 and

Table 1: Results for Log-Rank Test for the LOS.

Variable	Chi-Squared	Df	Sig
Gender	89.926	1	< 0.0001
Age	0.035	1	0.852
Chronic	406.734	1	< 0.0001
Surgery	579.922	1	< 0.0001
Emergency	151.917	1	< 0.0001
Comorbidity	509.641	6	< 0.0001
Trauma	319.087	1	< 0.0001
Discharge	2016.433	3	< 0.0001
Principal Diagnostic Group	789.523	4	< 0.0001
Ward	18914.696	196	< 0.0001

the loglikelihood for 6 phases was -458860, with the Chi-squared test for nested models having a $p\text{-value}$ =0.652. With reference to the latter result, it should be noticed that a stage is nested in the subsequent one and the chi-square can be applied accordingly (Marshall and McLean, 2004; Garg, McClean, Meenan and Millard, 2011). However, it is important to note that some of the parameters μ_i associated with phase 2 and 3 in Table 2 are equal to zero. This would suggest that no one is observed leaving the first phase and third phase. These are aggregated with the neighbouring phase, in this case phase 2 to produce a three stage model. Hereafter, the term stage will indicate a set of sequential phases with estimates of μ_i approaching to zero aggregated together with the closest phase associated to a strictly positive μ_i as indicated in equation 7. In effect, the phases with small values of μ_i parameters are redundant and only the most dominant phases with the largest μ_i values are meaningful. This will also prevent an over-fitted model as reported in earlier literature (Marshall et al. 2012). In each stage we calculated the probability of leaving the hospital ward due to death, transfer or returning home.

Table 3 reports the patients leaving the ward, the interval LOS and the average LOS for

Table 2: Results of fitting the Coxian phase-type distribution.

Phase i	$\frac{\hat{\lambda}_i}{\hat{\lambda}_i + \hat{\mu}_i}$	$\frac{\hat{\mu}_i}{\hat{\lambda}_i + \hat{\mu}_i}$	π_i
1	1	0	0
2	0.934	0.066	0.066
3	1	0	0
4	0.213	0.787	0.735
5	-	-	0.199

each stage. For example, 16386 patients leave during the first stage within 3 days and their duration lasts on average 2 days. The first stage in the length of distribution (0-3 days) can be considered the shorter stay patients. The second stage (4-17 days) would be considered those patients who have a much longer duration of stay in hospital. The third and final stage (> 17 days) is regarded as the patients with a longer LOS in hospital.

Table 3: Average and interval LOS at each stage.

Stage	Num. of patients (leaving at the stage)	Lower bound LoS	Upper bound LoS	Average LoS
1	16386	0	3	1.98
2	100244	4	17	9.19
3	21567	18	100	26.43

4.2 THE ANALYSIS OF THE IN-HOSPITAL RATE OF MORTALITY: INCLUDING THE COXIAN STAGES AS A CONTRIBUTION TO THE MULTILEVEL MODEL

In this section we evaluate the Italian geriatric wards for the in-hospital rate of mortality. The crude in-hospital death rate was on average 10.5 with a minimum value equal to 0 and

a maximum value equal to 34 per 100 discharges. In-mortality death rate was the response of a logistic regression model (8), whose intercept was considered as being random at both the hospital and regional levels. The covariates considered were

- at the patient level: Gender, Age, **LOS**, Admission in Emergency, Comorbidity (Elixhauser Index), Surgery, Trauma, Chronic disease, Disease (Circulatory, respiratory, other);
- at the ward level: Number of beds and Ownership (In this case, the Hospital's ownership was considered as a ward level);
- at the regional level: Regional Index of Good Health and Regional political government.

We wish to explore the advantage of results obtained in the analysis of **LOS**. In particular, we will change the continuous covariate **LOS** (in days) to the categorical variable *Stage*. In general the replacement of the continuous variable with a related categorical variable could give a better interpretation of the odds ratio for a logit model. We will evaluate the improvement on the model using the AIC value. In particular the multilevel model with the complete set of covariates has a value of AIC equal to 85296, while the same model with the categorical covariate *Stage* has a value of AIC equals to 79438. This result suggests that the inclusion of the stages of the Coxian phase-type distribution has a significant contribution in the multilevel logit model to evaluate the effectiveness among the Italian geriatric wards. The table 4 reports the results regarding the multilevel logistic models.

First of all we tested the null hypothesis that the full model is equal to the model without covariates. The Chi-Squared test for nested model leads to rejecting the null hypothesis with a *p-value* less than 0.0001. The Intraclass Correlation Coefficient (ICC), defined as the proportion of variance that is accounted for by the group level, is reported in table 4 (and is indicated as ρ). The ICC of the empty and complete models are statistically different, denoting the validity of the multilevel approach. Moreover when patient, ward and region characteristics were added to the model (complete model), the ICC significantly decreased. Model *B* shows that mortality was significantly affected by the patients' age, gender and

Table 4: Results for multilevel logistic models. (In table: ***=p-value< 0.0001, **=p-value< 0.001, *=p-value< 0.01)

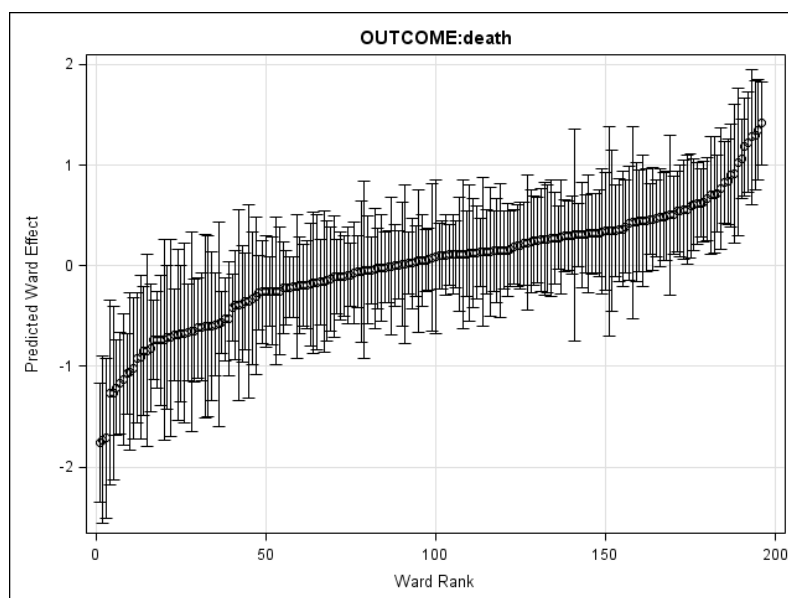
	Model A (empty)		Model B (complete)		
	Estimate	SE	Estimate	SE	
Fixed Effects					
Constant	-2.416	0.1647	-1.449	0.136	***
Female			0.185	0.018	***
> 85 year			0.522	0.019	***
Stage 1			1.457	0.030	***
Stage 2			-0.455	0.025	***
Emergency			0.530	0.041	***
Comorbidity			-0.051	0.010	***
Surgery			0.254	0.030	***
Trauma			0.141	0.045	**
Chronic Disease			-0.131	0.022	***
Circulatory/Respiratory Disease			0.235	0.020	**
Beds number			-0.004	0.002	
Ownership: public			0.557	0.102	***
Index of good health			-0.012	0.006	***
Political government: democrats			-0.522	0.019	**
Random Effects					
u_{0jk}	0.540	0.053	0.452	0.047	
γ_{0k}	0.517	0.203	0.268	0.113	
ρ ward	0.124		0.113		
ρ region	0.119		0.067		
Deviance	87239.414		79408.173		

emergency admission. This indicates that older patients, female and those patients with emergency admission had higher risk of dying. Multiple co-morbidities have a lower risk of dying than patients with fewer reported co-morbidities. Patients with a surgical diagnosis or those admitted with a trauma are at greater risk of dying than patients diagnosed with a medical condition or admitted without trauma. The incorporation of the stage covariate also seems appropriate. Patients leaving during the first stage are at greater risk of dying than patients leaving in the third stage while patients leaving during the second stage are at lower risk of dying than patients leaving on the third stage. Patients with a chronic disease are at lower risk of dying than patients diagnosed with a non chronic condition. Circulatory and respiratory diseases lead to a higher risk of dying than other diseases. By considering the ward characteristics, public hospitals had a higher risk of mortality compared to the other, while the number of beds was not significantly associated with the risk of dying. At the regional level, the index of good health seems to affect negatively the risk of mortality, while the patients in democratic regions has a lower risk of dying than those in the republic regions. Figure 4 shows the rank of the wards with respect to the mortality rate, according to the value of u_{0jk} . As indicated in Section 2.2, the 95% confidence intervals identify wards under or over the mean (that is represented by the value 0 on the co-ordinates) of risk of mortality. In particular a 95% confidence interval about the mean highlights a ward with a risk of mortality greater than the overall mean. On the Figure 4, adjusted for the other covariates, 33 wards have a risk of mortality greater than the overall mean.

5. CONCLUSION

This paper introduces a multi-level model that incorporates information taken from the Coxian phase-type distribution so to represent a survival variable. The simple approach is illustrated with an example from healthcare by modelling patient in-hospital mortality in

Figure 4: The rank of the geriatric wards respect to the mortality rate



geriatric wards in Italy. The multi-level model is used to provide a means of controlling for the existence of possible intra-ward and intra-hospital correlations. The Coxian phase-type distribution is used to represent the LOS (survival) in stages and used as a covariate in the multi-level model. The resulting multi-level model with covariate stage provides a better representation of the data and enhances that obtained by just having the multi-level model on its own. By having this more accurate model, better assessments can be made of the different hospital wards and regions and their care and cost for elderly patients staying in hospital.

This work illustrates the merit of the inclusion of the Coxian phase-type distribution into the multi-level model. Further research will attempt to mathematically combine the two approaches and to evaluate the new methodology by performing a comparison with Faddy et al. (2009) and McGrory et al. (2009)'s methods where a Coxian phase-type model with covariates is used to predict class membership.

BIBLIOGRAPHY

Berta P., Seghieri C., Vittadini G., (2013) Comparing health outcomes among hospitals: the experience of the Lombardy Region. *Health Care Manag Sci*, 16:245-257 DOI 10.1007/s10729-013-9227-1

Donabedian A. (2005) Evaluating the quality of medical care. *Milbank Q* 83:691729. *The Journal of the American Medical Association*, 260(12), 1743-1748.

DesHarnais S., McMahon L.F., Jr, Wroblewski R. (1991) Measuring outcomes of hospital care using multiple risk-adjusted indexes. *Health Serv Res*;26(4):425-445.

Ebbes, P., Bockenholt, U. (2004) Regressor and random-effects dependencies in multilevel models, *Statistica Neerlandica* Volume 58, Issue 2, 161-178, May

Ellis, G., Whitehead, M. A., O' Neil, D.L., Langhorne, P., (2011) Comprehensive geriatric assessment for older adults admitted to hospital: meta-analysis of randomised controlled trials. *BMJ* 2011;343:d6553 doi: 10.1136/bmj.d6553

Faddy, M., 1994. Examples of fitting structured phase-type distributions. *Applied stochastic models and data analysis*, 10, 247-255.

Faddy M.J. and McClean S.I. (1999) Analysing Data on Lengths of Stay of Hospital Patients Using Phase-type Distributions, *Applied Stochastic Models in Business and Industry* 15, 311-317.

Faddy M.J., McClean S.I. (2005) Markov chain modelling for geriatric patient care. *Methods Inf Med*, 44, 369-373.

Faddy M, Graves N, Pettitt (2009) A. Modeling Length of Stay in Hospital and Other Right Skewed Data: Comparison of Phase-Type, Gamma and Log-Normal Distributions. *Value in Health*, 12, 309-314.

Garg, L., McClean, S., Meenan, B. J., Millard, P. (2011) Phase-Type Survival Trees and Mixed Distribution Survival Trees for Clustering Patients Hospital Length of Stay, *Informatica*, Vol. 22, No. 1, 57-72

Goldstein H. (1995) Multilevel statistical models. Arnold, London

Hox J.J. (1995) Applied multilevel analysis. TT-Publikaties, Amsterdam

Jones, K., Wright, C., Bell, A. (2012) Do multilevel models ever give different results?. In preparation, <https://www.researchgate.net/publication/252146040>

Jones, K., Subramanian, K. S. V. (2012, updated 2014) Developing multilevel models for analysing contextuality, heterogeneity and change Volumes 1, 2. University of Bristol.

Kroch E. A., Duan M., Sillow-Carrol S., Meyer J.A. (2007) Hospital performance improvement: trends in quality and efficiency. A quantitative analysis of performance improvement in U.S. hospitals. The Commonwealth fund.

Institute of Medicine Medicare (1990) A Strategy for Quality Assurance. Volume I. K. N. Lohr, ed. Washington, DC: National Academy Press.

Lau, T., Fang, C. and Leung, F. (2013) The Effectiveness of a Geriatric Hip Fracture Clinical Pathway in Reducing Hospital and Rehabilitation Length of Stay and Improving Short-Term Mortality Rates Geriatric Orthopaedic Surgery and Rehabilitation vol. 4 no. 1, 3-9.

Librero, J., Peirò, S., Ordina, R. (1999) Chronic Comorbidity and Outcomes of Hospital Care: Length of Stay, Mortality, and Readmission at 30 and 365 Days *Clin Epidemiol* Vol. 52, No. 3, pp. 171-179.

Marshall A.H., McClean S.I. (2003) Conditional phase-type distributions for modelling patient length of stay in hospital. *Inter. Trans. Op. Res.*, 10, 565-576.

Marshall A.H., McClean S.I. (2004) Using Coxian Phase-Type Distributions to Identify

Patient Characteristics for Duration of Stay in Hospital. Health Care Management Science 7, 285-289

Marshall A.H., Zenga M. (2012) Experimenting with the Coxian Phase-Type distribution to Uncover Suitable Fits, Methodology and computing in Applied probability, 14, pp. 71-86, <http://dx.doi.org/10.1007/s11009-010-9174-y>

McGrory CA, Pettitt AN, Faddy MJ. (2009) A fully Bayesian approach to inference for Coxian phase-type distributions with covariate dependent mean. Computational Statistics and Data Analysis, 53, 4311-4321.

Millard, P.M. (1991) Throughput in a department of geriatric medicine: a problem of time, space and behaviour. Health Trends, 24, 20-24.

Neuts M.F. (1975) Computational Uses of the Method of Phases in the Theory of Queues, Computers and Math. with Appl., 1, 156-66.

Neuts M.F. (1981) Matrix-Geometric Solutions in Stochastic Models - An Algorithmic Approach, John Hopkins University Press.

OECD (2004) Ageing societies and the looming pension crisis. Paris: OECD

Rice N, Leyland A (1996) Multilevel models: applications to health data. J Health Serv Res Policy 1,154-164

Spencer, N. H. and A. Fielding (1998), An instrumental variable consistent estimation procedure to overcome the problem of endogenous variables in multilevel models, University of Hertfordshire.

Spencer, N. H. and A. Fielding (2002), A comparison of modelling strategies for value-added analyses of Educational Data. Computational Statistics, Vol. 17, No. 1, 2002, p. 103-116.

Spiegelhalter, D. J. (1998), Bayesian graphical modelling: a case-study in monitoring health outcomes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47: 115133.
doi: 10.1111/1467-9876.00101